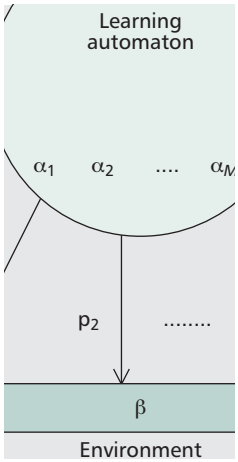
ACCEPTED FROM OPEN CALL

ADAPTIVE WIRELESS NETWORKS USING LEARNING AUTOMATA

PETROS NICOPOLITIDIS, GEORGIOS I. PAPADIMITRIOU, AND ANDREAS S. POMPORTSIS,
ARISTOTLE UNIVERSITY OF THESSALONIKI
PANAGIOTIS SARIGIANNIDIS, UNIVERSITY OF WESTERN MACEDONIA
MOHAMMAD S. OBAIDAT, MONMOUTH UNIVERSITY



Learning automata are artificial intelligence tools used in areas where adaptivity to the characteristics of the wireless environment can result in a significant increase in network performance. The authors review current approaches in using learning automata to provide adaptivity to wireless networking.

ABSTRACT

Wireless networks operate in environments with unknown and time-varying characteristics. The changing nature of many of these characteristics will significantly affect network performance. This fact has a profound impact on the design of efficient protocols for wireless networks and as a result adaptivity arises as one of the most important properties of these protocols. Learning automata are artificial intelligence tools that have been used in many areas where adaptivity to the characteristics of the wireless environment can result in a significant increase in network performance. This article reviews state of the art approaches in using learning automata to provide adaptivity to wireless networking.

INTRODUCTION

Wireless networks operate within a dynamic environment, which entails possibly unknown and time-varying characteristics, with the most common ones being the following:

Variable link qualities: These are primarily caused by multipath fading and co-channel interference. The time variability in the quality of a link leads to the need for adaptive operation of several protocols across the protocol stack. For example, at the physical layer, an increased bit error rate (BER) should be countermeasured by either using a more robust modulation scheme or an increase in transmission power via power control. At the transport layer, the TCP congestion window (cwnd) should be adaptively handled so as not to misinterpret transmission losses for congestion.

Dynamic topologies: In most wireless networks nodes are typically mobile and have a fixed transmission range. Thus, the topology of the network will change with time and the network nodes should adapt to such changes. Typical examples of protocols that need to adapt to dynamic topologies are wireless medium access control (MAC) and routing protocols.

Power management: The mobile nodes of wireless networks are typically battery-powered. Therefore, specific measures have to be taken in the direction of adapting the energy consumption of a node to its residual energy. Power control mechanisms that are typically employed in wireless networks must also take this need into account.

Spectrum usage: Due to the fact that the spectrum has become overcrowded, the research trend is toward wireless cognitive systems that can opportunistically use licensed parts of the spectrum when these are not in use. This implies profound challenges for several protocol layers.

Changing traffic patterns: The needs of nodes for medium access can change over time according to the needs of the applications being served, thus leading to the need for adaptive channel access protocols.

Learning automata (LA) [1] have been found to be useful in systems that possess incomplete knowledge regarding the environment in which they operate. In the area of data networking, LA have been applied to several problems, including the design of self-adaptive MAC, routing, and transport layer protocols. This article surveys state-of-the-art approaches in using LA to enhance the performance of wireless networks. After providing introductory coverage of LA, we discuss state-of-the-art research on using LA at several layers of the wireless networking stack and in specific wireless networking environments, such as sensor networks and wireless data broadcasting networks.

REINFORCEMENT LEARNING AND LEARNING AUTOMATA

In the area of learning systems, reinforcement learning has emerged as a promising technique. Reinforcement learning aims to provide a system with the necessary information in order to plan its actions so as to maximize the reward it receives from the environment in which it operates. Reinforcement learning techniques typically utilize the following triple:

In the area of learning systems, reinforcement learning has emerged as a promising technique. Reinforcement learning aims to provide a system with the necessary information in order to plan its actions so as to maximize the reward it receives from the environment in which it operates.

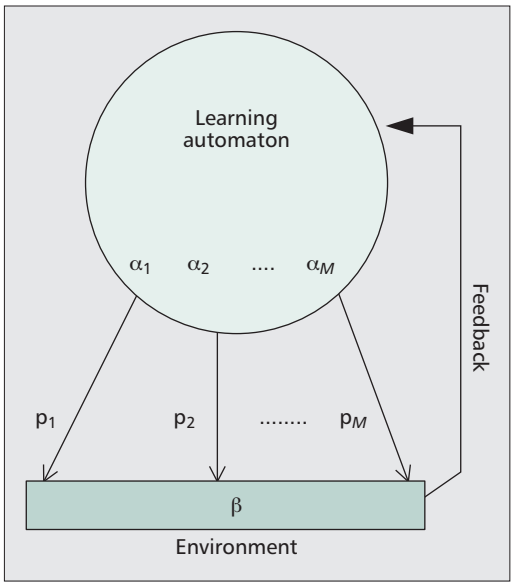


Figure 1. *Operation of a learning automaton.*

- The number of environmental states S
- The number of possible actions M to be taken by the system
- An environmental reward β for each action taken by the operating system

At each time instant during its operation, a system based on reinforcement learning that resides in a certain state chooses one of the available actions, performs it, and receives a new state from the environment as well as the environmental response. By repeating the above procedure, the goal of a reinforcement learning system is to achieve maximization of the received environmental reward.

Learning automata are artificial intelligence tools whose operation can be viewed in the framework of reinforcement learning. The operation of an LA is shown schematically in Fig. 1. An LA continuously interacts with the random operating environment so as to find among a set of actions the one that minimizes the average penalty the system receives by the environment. To achieve this, an LA uses a vector $p(n)$, which maintains the probability distribution for choosing at cycle n action $\alpha(n)$ from the set of actions $\alpha_1, \alpha_2 \ldots, \alpha_M$. Obviously,

$$\sum_{i=1}^{M} p_i(n) = 1.$$

An LA can be used only if there is a feedback mechanism that conveys to the LA the environmental response to each performed action.

The operation of an LA is based on the probability updating algorithm, also known as the reinforcement scheme. This algorithm uses the environmental response that was received as a result of performing action a_i selected at cycle n (action $\alpha(n)$) in order to update the probability distribution vector p. After the updating is performed, the LA selects the action to perform at cycle $n + 1$, according to the updated probability distribution vector $p(n + 1)$. A general reinforcement scheme has the form

$$\begin{aligned} p_i(n+1) &= p_i(n) - (1-\beta(n))g_i(p(n)) \\ &\quad + \beta(n)h_i(p(n)), \ \forall\, a(n) \neq a_i \\ p_i(n+1) &= p_i(n) + (1-\beta(n))\sum_{j\neq i} g_j(p(n)) \\ &\quad - \beta(n)\sum_{j\neq i} h_j(p(n)), \ \text{if } a(n) = a_i. \end{aligned} \tag{1}$$

The functions g_i and h_i are associated with reward and penalty for the selected action a_i, respectively, while $\beta(n)$ is a parameter expressing the received environmental response at cycle n, normalized in the interval [0,1]. The lower the value of $\beta(n)$, the more favorable the response.

If the environmental response is of a binary nature, indicating only reward or penalty via 0 or 1, respectively, the environment is known as P-model, and the LA is known to operate within a P-model environment. Nevertheless, as in many cases a P-model LA will yield only a gross estimate of the environmental response, other schemes have also appeared. In these, the environmental response can take values in [0..1], thus indicating actions that are neither completely rewarding nor penalizing. Specifically, in a Q-model LA, the environmental response can take more than two (still finite, however) possible values in the interval [0..1], whereas in an S-model LA, the environmental response can take continuous values in [0..1].

Regarding the choices made for functions g_i and h_i, different selections for these functions result in a number of different reinforcement schemes, the most common being the following.

The Linear Reward-Penalty (LR-P) Scheme — In this scheme, after selection of action $a(n)$ at cycle n, the following reinforcement scheme is applied:

$$\begin{aligned} p_i(n+1) &= p_i(n) + \beta(n) \\ &\quad [L/(M-1) - L\,(p_i(n) - a)] \\ &\quad - [1 - \beta(n)]\, L\,(p_i(n) - a), \ \forall\, a(n) \neq a_i. \end{aligned} \tag{2}$$

$$\begin{aligned} p_i(n+1) &= p_i(n) - \beta(n)\, \mathrm{L}\,(p_i(n) - a) \\ &\quad + [1 - \beta(n)]\, L\, [1 - (p_i(n) - a)], \\ &\quad \text{if } a(n) = a_i. \end{aligned}$$

In the above equation pair, M is the number of actions, and L is a parameter that lies in (0..1) and defines the learning rate of the LA. In this scheme, g_i and h_i are linear functions of the corresponding action probabilities p_i. Thus, in a P-model LR-P ($P_{LR\text{-}P}$) automaton, after reception of a favourable response for the action that was selected in the previous cycle, the corresponding probability of selecting this action again is increased. After reception of an unfavorable response, however, the probability of selecting this action again is decreased. For the Q and S model LR-P schemes, probability updates occur as combinations of g_i and h_i weighed by $1 - \beta(n)$ and $\beta(n)$, respectively.

The Linear Reward-Inaction (LR-I) Scheme — In this scheme, after selection of action $a(n)$ at cycle n, the following reinforcement scheme is applied:

$$p_i(n+1) = p_i(n) - L(1-\beta(n)) \\ (p_i(n) - a), \forall a(n) \neq a_i. \\ p_i(n+1) = p_i(n) + L(1-\beta(n)) \\ \sum_{j \neq i} (p_j(n) - a) \text{ if } a(n) = a_i. \quad (3)$$

In this scheme, g_i is a linear function of p_i, and h_i always equals 0. Thus, in a P-model LR-I ($P_{LR\text{-}I}$) automaton, after reception of a favorable response ($\beta(n) = 0$) for the action that was selected in the previous cycle, the corresponding probability of selecting this action again is increased. When an unfavorable response ($\beta(n) = 1$) is received, however, the probability of selecting this action again is not decreased, but remains the same. For the Q and S model LR-I schemes, probability updates occur as functions of g_is only.

Nonlinear Schemes — In this case g_i and h_i are nonlinear functions of p_i.

For the parameters L and a in Eqs. 2 and 3, it holds that L and a take values in (0,1) and p_i takes values in (a,1). L defines the speed of the automaton convergence. The lower the value of L, the more accurate the estimation made by the automaton, a choice, however, that comes at the expense of convergence speed. Parameter a is used to enhance the adaptivity of the LA. This is because when the choice probability of an action approaches zero, this action is not selected for a long period of time. However, after a time period the changing nature of the environment might render this action a favorable one. Thus, since this action now has zero probability of being selected, the LA will not be capable of selecting it anymore. Thus, the use of a non-zero value for parameter a prevents the choice probabilities of actions from taking values of zero and thus increases the adaptivity of the LA.

To better understand parameters L and a, Figs. 2 and 3 show their effect on the convergence of a $P_{LR\text{-}P}$ LA. In this experiment, we assume an action a_i with an actual probability $d_i = 0.8$ of being the optimal one, with this probability changing to a new one (dashed line, $d_i = 0.8$) after about 30,000 cycles and we plot the probability estimate of the LA for this action. One can see that the choice for the value of L reflects the classic speed versus accuracy problem. As can be seen from Fig. 2, where $L = 3*10^{-2}$, large values of L (compared to that used in Fig. 3) provide a higher convergence speed at the expense, however, of convergence accuracy. Contrary, with a small value of L (Fig. 3), we get better convergence, at the expense, however, of convergence speed. Finally, convergence accuracy also depends on the value of a, with smaller values of a giving better convergence. Increased values of a will make the estimated probability p_i converge to a point higher than d_i. Figure 3 supports this fact.

By altering the selection procedure for the action to be rewarded or penalized after each cycle, more efficient LA in terms of convergence speed can be derived. This idea is implemented in pursuit LA, which employ a vector that contains the running estimates of the environmental response for each action and at each probability update the LA will always reward the action with the current minimum penalty estimate, thus always pursuing the action with the highest reward.

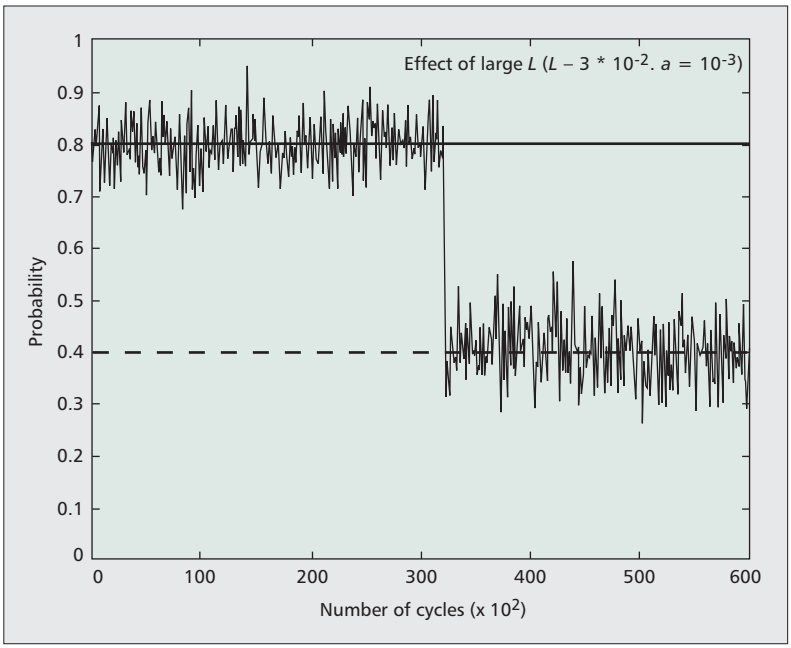


Figure 2. *Effects of a large value of the learning rate* L *on* $P_{LR\text{-}P}$ *LA convergence.*

Applications of LA to Wireless Networking

Applications to Different Layers of the Network

Physical Layer — LA have recently been applied to determine the transmission power of mobile nodes. In [2], the authors model the distributed power control problem in an infrastructure wireless network as a non-zero sum game between the mobile nodes. Two LA-based algorithms for distributive solution of the power control game are proposed. In both of them, each node operates an LA, which determines the probability of choosing a certain transmission power based on the feedback received from the base station (BS). This feedback is essentially the environmental response for choosing a certain power level and expresses the amount of information that a mobile node can transfer during the lifetime of its battery. However, the proposed algorithms can also operate for other definitions of the environmental response as well.

LA have also been applied to dynamically adapt the transmission rate in wireless links in a physical/MAC layer synergy. In [3], the authors present a joint user scheduling and adaptive rate control for downlink wireless transmission that can be easily applied to third-generation (3G) wireless systems. User scheduling is performed at the MAC layer and is enabled via an LR-P LA, while the rate selection procedure is implemented in the physical layer via a pursuit reward-inaction LA. Via utilizing information from acknowledg-

ment packets for downlink data, the joint algorithm is shown to be able to learn and use the best transmission rate according to the conditions of the channel. Simulation results show that the proposed algorithm is suitable for low-mobility applications in 3G wireless networks.

In [4], an LA-based scheme, the Stochastic Automata Rate Adaptation Algorithm (SARA), is proposed in order to achieve rate adaptation. According to it, each transmitter has a list of possible modulation schemes to use, each yielding a different transmission rate. The selection of the rates to use and the probability of using each one of these rates is dynamically updated via an LA based on the obtained feedback from the receiver. SARA, which is fully compatible with the IEEE 802.11 MAC protocol, is compared via simulation under different channel scenarios to Automatic Rate Fallback Adaptive Automatic Rate Fallback and the scheme presented in [3]. The results show superior performance of SARA compared to the previous approaches and a lower computational complexity for SARA for achieving rate adaptation compared to the scheme presented in [3].

MAC Layer — In the context of medium access, LA have found use in both infrastructure and ad hoc wireless LANs (WLANs). In [5, 6] the ability of LA to learn the parameters of the operating environment is exploited in order to provide efficient MAC protocols for bursty traffic WLANs. It is proved for both the approaches in [5, 6] that the learning algorithm asymptotically tends to assign to each node a portion of the bandwidth proportional to the node's needs.

In [5], the BS is equipped with a $P_{LR\text{-}P}$ LA that maintains the probability of granting permission to transmit to each of the mobile nodes under its coverage. The BS polls the mobile nodes according to these estimates, and after each poll, the network feedback information is used in order to update the choice probability of each mobile node. Since the offered traffic is of bursty nature, when the BS realizes that the selected node had a packet to transmit, it is probable that the selected node will also have packets to transmit in the near future. Thus, its choice probability is increased. On the other hand, if the selected node notifies that it does not have buffered packets, its choice probability is reduced, since it is likely to remain in this state in the near future. The same idea applies to the condition of bursty error wireless links. Thus, when the BS fails to receive feedback about the selected mobile's state, the mobile is probably experiencing a relatively high BER link to the BS. Thus, the choice probability of the selected node is lowered in order to reduce the chance of a future futile poll. This polling protocol, named Learning Automata Based Polling (LEAP), is compared to the Randomly Addressed Polling (RAP) and Group Randomly Addressed Polling (GRAP) polling protocols for WLANs via simulation and is shown to exhibit superior performance under bursty offered traffic. Some of these results are regenerated in Fig. 4, which reveals the performance superiority of LEAP in terms of throughput, a result attributed to its capability of adapting to the bursty offered traffic.

In [6], the same approach is taken for an ad hoc WLAN; in this case every mobile node is equipped with a $P_{LR\text{-}P}$ LA. The proposed protocol, named Ad Hoc Learning-Automata-Based Protocol (AHLAP), is compared to IEEE 802.11 distributed coordination function (DCF) under bursty traffic conditions via simulation. Some of the simulation results are shown in Fig. 5 for low- and high-grade burstiness of the offered traffic. One can easily see that the burstier the network traffic, the better the behavior of AHLAP against IEEE 802.11 DCF; a fact attributed to the ability of AHLAP to pinpoint the active nodes in the network and grant them permission to transmit.

The variability of bursty offered traffic is addressed from another perspective in [7], where the authors propose a MAC protocol for clustered wireless ad hoc networks. Inside each cluster, medium access is controlled by a designated clusterhead node via a time-division multiple access (TDMA) scheme and as a result intra-cluster transmissions are collision-free. Inter-cluster communications are served by a code-division multiple access (CDMA) scheme overlaid on the TDMA technique so as to achieve interference-free communication between nodes in different clusters. In order for the above scheme to work, the network needs cluster formation, code assignment and slot assignment algorithms, all three of which are based on LR-P LA. For cluster formation, each node operates an LR-P LA, which results in organization of the nodes in a way that produces a minimum number of non-overlapping clusters. The code assignment algorithm is implemented using an LR-P LA at each clusterhead to achieve spatial reuse of the limited number of CDMA codes that can enable concurrent intercluster communication. Finally, for slot assignment to nodes, an LA at each clus-

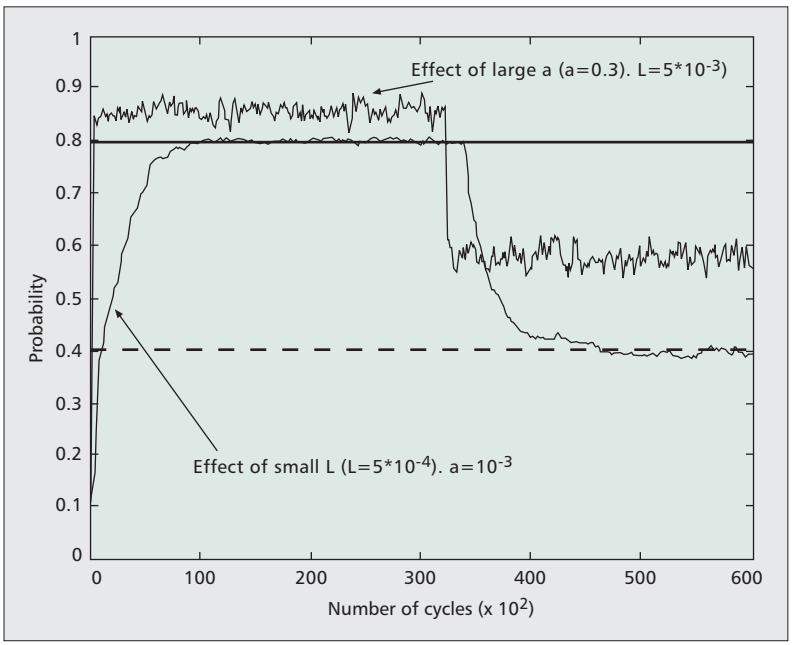


Figure 3. *Effects of a small value of the learning rate* L *and of a large value of parameter* a *on* $P_{LR\text{-}P}$ *LA convergence.*

terhead is used to assign to each node of the cluster a fraction of the TDMA frame proportional to the traffic load of the node. Simulation results in [7] show that the proposed scheme outperforms existing methods under bursty traffic conditions.

Network Layer — LA have also been used for multicast routing in wireless ad hoc networks [8]. Based on predictions of node mobility, the work in [8] finds the routes with higher lifetimes, based on which a stochastic graph representing the virtual multicast backbone of the ad hoc network is built. Then a distributed LA-based algorithm is applied to this graph to solve the multicast routing problem. Simulation results in [8] reveal that it outperforms existing algorithms.

Multicast routing is also addressed in [9], whose contribution is twofold: first, it proposes three LA-based algorithms that find the optimal solution to the minimum weighted Steiner connected dominating set problem; second, one of these algorithms is implemented in a distributed manner in the network nodes to solve the multicast routing problem. Simulation results reveal the superiority of the proposed approach against well-known multicast routing protocols under a variety of performance metrics.

The virtual backbone graph mentioned in the context of [8] above is also exploited in [10] to solve via global flooding the broadcast storm problem that is inherent in broadcast-based routing. A set of LA operating on the network nodes is used to route traffic over the virtual backbone via broadcast routing. This results in mitigation of the broadcast storm problem as the number of nodes implementing the broadcasts matches the number of nodes in the backbone. Simulation results in [10] reveal that the proposed algorithm enjoys significantly higher performance than that of similar existing algorithms with only a minor increase in the overhead of exchanged control messages.

Transport Layer — The operation of TCP over the error-prone wireless links can lead to significant degradation of its performance, due to the fact that TCP cannot differentiate congestion losses from reception losses. To combat this, LA are used in [11] in order to recognize the two kinds of losses by observing the arrival of acknowledgment and duplicate acknowledgment packets. Thus, the proposed protocol will dynamically adapt to the changing network conditions and appropriately update the cwnd size. Simulation results in [11] under varying wireless network conditions show up to 18 percent throughput increase and 55 percent packet loss reduction compared to the corresponding figures of traditional TCP.

Another transport layer issue, congestion avoidance, has recently been treated with LA in the context of sensor networks for healthcare applications [12]. In this approach, at each intermediate node in the path from the source to the destination node, an LA continuously interacts with the environment. Based on the offered traffic at each node, each such LA adaptively learns the optimal data rate that should be flowing through the respective node so that congestion is locally controlled at each node. Simulation results in [12] show that the proposed algorithm can efficiently avoid congestion in typical healthcare wireless sensor networks.

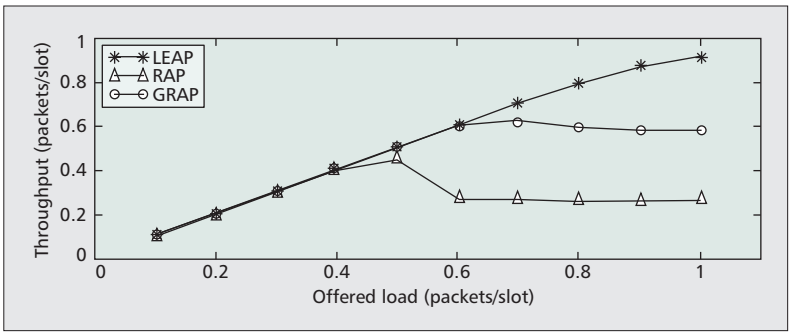


Figure 4. *Throughput of LEAP, RAP, and GRAP vs. the aggregate normalized network offered load.*

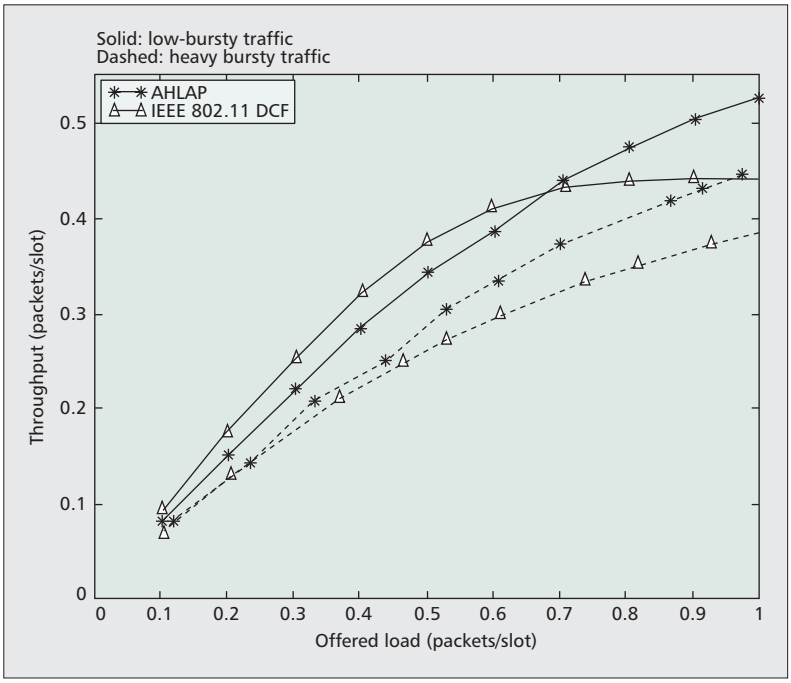


Figure 5. *Throughput of AHLAP and IEEE 802.11 DCF, for low- and high-grade bursty traffic, vs. the aggregate normalized offered load to the network.*

Applications to Areas of Wireless Networks

Applications to Ad Hoc Networks — In [13] an LA-based approach is used to combat the problem of performance degradation that occurs as the size of the network grows. After proposing an LA-based centralized algorithm for finding the near optimal solution for the minimum connected dominating set representation of the network, the authors propose a distributed version of the centralized algorithm that can be implemented at every network node. Simulation results reveal the superiority of the proposed approach over the existing ones.

Applications to Sensor Networks — The primary concern of wireless sensor networks is to reduce energy consumption. This can be achieved by data aggregation, which is performed in intermediate nodes en route to the sink. Data aggrega-

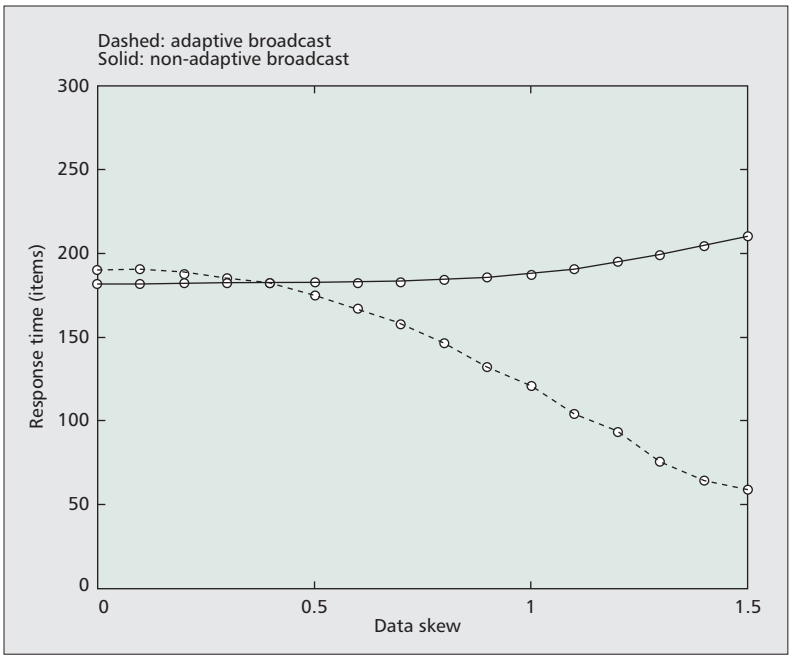


Figure 6. *Performance gains of the method of [19] due to the incorporation of adaptivity.*

tion prevents redundant information being forwarded by neighboring nodes that sense the same event. The benefits of aggregation are maximized when the network routes data packets over routes comprising nodes with a high probability of having sensed the same information as that carried in the routed packet. Since these paths are not static, but rather follow changes in the environment, [14] proposes an aggregation scheme where each node is equipped with an LA in order to collectively learn the path of aggregation with maximum aggregation ratio. Simulation results in [14] show that the proposed method exhibits significantly better performance than other aggregation mechanisms, especially in dynamic environments.

LA have also been used in [15]. This work proposes a fully distributed LA-based algorithm that efficiently clusters wireless sensor networks in order to reduce energy consumption and increase network lifetime. Contrary to other approaches which recluster the network at regular time intervals, in the proposed algorithm reclustering is performed locally when network conditions mandate it, thus resulting in a decrease in energy consumption. Simulation results in [15] reveal increased network lifetime and energy efficiency compared to other clustering methods.

In [16] the authors propose an LA-based scheduling algorithm to deal with the problem of using the fewest possible number of sensors to detect moving objects via a sensor network. This objective of course translates to reduced energy consumption for the network. According to the proposed approach, each sensor node is equipped with a set of LA. The objective of each such set is to learn the maximum sleep duration for the node so that the detection rate of targets by the node does not fall dramatically. Simulation results in [16] reveal increased energy efficiency over existing methods.

Cognitive Radio — LA can also provide a powerful learning tool in the environment of cognitive wireless networks. Reference [17] studies the problem of quality of service (QoS) enabled opportunistic spectrum access for non-licensed users. For the case when not all licensed users can be supported over the available spectrum, each node should operate an LA. A set of priorities is defined and the probability that a node can access the medium is set to be proportional to the node's priority. Then an LA-based algorithm is used to solve the priority-based medium access in a distributed way.

In [18], the authors propose a method of managing the network routes that carry traffic between non-licensed users in a cognitive multihop network. The problem addressed is attributed to the randomness derived from unpredictable activity of licensed users. In the proposed approach, LA are used at the network nodes in order to find those routes least affected by licensed users.

Wireless Data Broadcasting — In the area of wireless data broadcasting, [19] proposed an LA-based push system that enables adaptivity in environments with dynamic and a priori unknown client demands. In such environments, there is a large number of clients with overlapping demands for the various data items that reside in the server's database, which are delivered to the clients via broadcasting. In the system presented in [19], an $S_{LR\text{-}I}$ LA is incorporated at the broadcast server, which uses simple feedback from the clients to obtain an estimate of client demands. Simulation results in [19] show that under dynamic environments, where client demands change over time, the incorporation of adaptivity significantly improves performance. Some of these results are depicted in Fig. 6 for a server broadcasting 300 data items and show the performance gains in terms of mean waiting time for an item (response time). These gains for the adaptive system increase for increasing values of the data skew coefficient, a parameter that reflects the amount of commonality in the demands of various clients.

Conclusions

Adaptivity to the various unknown and time varying aspects of the wireless environment is a challenging research topic. Learning automata are artificial intelligence tools that have been used toward achieving the goal of such adaptivity. After discussing the dynamic properties of the wireless environment, this article overviews the structure, operation, and key parameters of LA as well as the impact of these parameters on their behavior. Next, it provides an up-to-date survey of state-of-the art approaches using LA to provide adaptivity to several layers of the wireless networking stack and to specific wireless networking environments, such as sensor networks and wireless data broadcasting. The goal of this survey is to summarize research in the area and bring forward the potential of LA as an important tool in the research toward intelligent wireless networking.

References

[1] M. S. Obaidat, G. I. Papadimitriou, and A. S. Pomportsis, "Learning Automata: Theory, Paradigms and Applications," *IEEE Trans. Sys., Man and Cybernetics-Part B*, vol. 32, no. 6, Dec. 2002, pp. 706–09.

[2] T. Joshi *et al.*, "SARA: Stochastic Automata Rate Adaptation for IEEE 802.11 Networks," *IEEE Trans. Parallel and Distrib. Sys.*, vol. 19, no. 10, Oct. 2008, pp. 157–90.

[3] M. A. Haleem and R. Chandramoulli, "Adaptive Downlink Scheduling and Rate Selection: A Cross-layer Design," *IEEE JSAC*, vol. 23, no. 6, June 2005, pp. 1297–97.

[4] Y. Xing and R. Chandramouli, "Stochastic Learning Solution for Distributed Discrete Power Control Game in Wireless Data Networks," *IEEE/ACM Trans. Net.*, vol. 16, no. 4, Aug.2008, pp. 932–44.

[5] P. Nicopolitidis, G. I. Papadimitriou and A. S. Pomportsis, "Learning-Automata-Based Polling Protocols for Wireless LANs," *IEEE Trans. Commun.*, vol. 51, no. 3, Mar. 2003, pp. 453–63.

[6] P. Nicopolitidis, G. I. Papadimitriou, and A. S. Pomportsis, "Distributed Protocols for Ad-Hoc Wireless LANs: A Learning-Automata-Based Approach," *Ad Hoc Networks*, Elsevier, vol. 2, no. 4, Oct. 2004, pp. 419–31.

[7] J. A. Torkestani and M. R. Meybodi, "An Efficient Cluster-Based CDMA/TDMA Scheme for Wireless Mobile Ad-Hoc Networks: A Learning Automata Approach," *J. Network and Comp. Apps.*, Elsevier, vol. 33, no. 4, July 2010, pp. 477–90.

[8] J. A. Torkestani and M.R. Meybodi, "Mobility-based Multicast Routing Algorithm for Wireless Mobile Ad-Hoc Networks: A Learning Automata Approach," *Comp. Commun.*, Elsevier, vol. 33, no. 6, Apr. 2010, pp. 721–35.

[9] J. A. Torkestani and M. R. Meybodi, "Weighted Steiner Connected Dominating Set and its Application to Multicast Routing in Wireless MANETs," *Wireless Pers. Commun.*, Springer, Feb. 2010.

[10] J. A. Torkestani and M. R. Meybodi, "An Intelligent Backbone Formation Algorithm for Wireless Ad hoc Networks based on Distributed Learning Automata," *Computer Networks*, Elsevier, vol. 54, no. 5, Apr. 2010, pp. 826–43.

[11] B. V. Ramana and C. S. R. Murthy, "Learning-TCP: A Novel Learning Automata Based Congestion Window Updating Mechanism for Ad Hoc Wireless Networks," *LNCS*, Springer, vol. 3769, 2005, pp. 454–64.

[12] S. Misra, V. Tiwari, and M. S. Obaidat, "LACAS: Learning Automata-based Congestion Avoidance Scheme for Healthcare Wireless Sensor Networks," *IEEE JSAC*, vol. 27, no. 4, May 2009, pp. 466–79.

[13] J. A. Torkestani and M. R. Meybodi, "Clustering the Wireless Ad-Hoc Networks: A Distributed Learning Automata Approach," *Parallel and Distrib. Comp.*, Elsevier, vol. 70, no. 4, Apr. 2010, pp. 394–405.

[14] M. Esnaashari and M. R. Meybodi, "Data Aggregation in Sensor Networks using Learning Automata," *Wireless Networks*, Springer, vol. 16, no. 3, Apr. 2010, pp. 687–99.

[15] M. Esnaashari and M. R. Meybodi, "A Cellular Learning Automata based Clustering Algorithm for Wireless Sensor Networks," *Sensor Letters*, vol. 6, no. 5, Oct. 2008, pp. 723–35.

[16] M.Esnaashari and M. R. Meybodi, "A Learning Automata based Scheduling Solution to the Dynamic Point Coverage Problem in Wireless Sensor Networks," *Computer Networks*, Elsevier, 2010.

[17] Y. Xing *et al.*, "Dynamic Spectrum Access with QoS and Interference Temperature Constraints," *IEEE Trans. Mobile Computing*, vol. 6, no. 4, Apr. 2007, pp. 423–33.

[18] Y. Song, C. Zhang, and Y. Fang, "Stochastic Traffic Engineering in Multihop Cognitive Wireless Mesh Networks," *IEEE Trans. Mobile Computing*, vol. 9, no. 3, Mar. 2009, pp. 305–16.

[19] P. Nicopolitidis, G. I. Papadimitriou, and A. S. Pomportsis, "Using Learning Automata for Adaptive Push-Based Data Broadcasting in Asymmetric Wireless Environments," *IEEE Trans. Vehic. Tech.*, vol. 51, no. 6, Nov. 2002, pp. 1652–60.

Biographies

PETROS NICOPOLITIDIS is an assistant professor in the Department of Informatics, Aristotle University of Thessaloniki, Greece. His research interests include wireless networks and mobile computing.

GEORGIOS PAPADIMITRIOU is an associate professor at the Department of Informatics, Aristotle University of Thessaloniki. His research interests include wireless networks, optical networks and learning automata.

PANAGIOTIS SARIGIANNIDIS is an adjunct lecturer at the University of Western Macedonia, Greece. His research interests include wireless networks, optical networks, and optical switching.

MOHAMMAD OBAIDAT [F'05] (obaidat@monmouth.edu) is a professor at Monmouth University, New Jersey. His research interests include wireless communications and networks, modeling and simulation, and security of computer and network systems. He is the Editor-in-Chief, Advisory Editor, and Editor of many international journals and transactions. He is the author of over 480 refereed journal and conference papers. He is the author of over 10 books. He has given numerous keynote speeches worldwide. He has received numerous awards. He is a Fellow of SCS.

ANDREAS POMPORTSIS is a professor at the Department of Informatics, Aristotle University of Thessaloniki. His research interests include computer networks and multimedia technology.

Adaptivity to the various unknown and time varying aspects of the wireless environment is a challenging research topic. Learning Automata are Artificial Intelligence tools that have been used toward achieving the goal of such adaptivity.